

MIKKO SALOILA

Comparison of Statistics Finland's and FSD's Metadata Models



FINNISH SOCIAL
SCIENCE DATA ARCHIVE

MIKKO SALOILA

Comparison of Statistics Finland's and FSD's Metadata Models

FINNISH SOCIAL SCIENCE DATA ARCHIVE 11, 2014
YHTEISKUNTATIEETEELLISEN TIETOARKISTON JULKAISUJA 11, 2014

Finnish Social Science Data Archive 11
Yhteiskuntatieteellisen tietöarkiston julkaisu 11
Mikko Saloila
2014
Comparison of Statistics Finland's and FSD's Metadata Models

Publisher: Finnish Social Science Data Archive
University of Tampere

Address: Finnish Social Science Data Archive
33014 University of Tampere

Tel: +358 40 190 1432
Sähköposti: fsd@uta.fi
Internet: <http://www.fsd.uta.fi>

ISSN 1459-8906
ISBN 978-951-44-9659-2

Graphic Design
Joona Ryödi, Redflow Inc.

Layout
Marita Alanko

Translation
Mirja Kytökari, InSight Translation and Communication Services
Tampere 2014

Contents:

1. Background and objectives	4
2 Metadata models: CoSSI and DDI Codebook	4
2.1 CoSSI	4
2.2 DDI Codebook	5
3 Data model comparison	5
3.1 CoSSI and DDI Codebook 2.1	5
3.2 CoSSI and DDI Codebook 2.5	6
3.3 DDI Lifecycle 3.1	6
3.4 Metadata test conversion	7
4. Public use files (PUFs)	8
4.1 Teaching datasets	8
4.2 Time-use data PUF	8
4.3 Census PUF	8
5. Publication of metadata in Statistics Finland's web service	9
6. Summary	9
7. Additional information	10
Annexes	11
Annex 1 Comparison table CoSSI-DDI Codebook (2.1) and Variable editor's data field descriptions	11
Annex 2 Comparison table CoSSI-DDI Codebook 2.5 (as per changed elements)	13
Annex 3 Comparison table for XML conversion	14

1. Background and objectives

The comparison of the metadata models of the Finnish Social Science Data Archive (FSD) and Statistics Finland was part of the *Meeting CESSDA-ERIC requirements* project funded by the Academy of Finland (decision no. 271853). One of its subprojects involved the development of metadata together with Statistics Finland.

Both Statistics Finland and FSD produce good-quality metadata for their research data. However, they each have their own metadata practices and models, due to which the metadata and the systems are not directly compatible. In other words, metadata produced by FSD is not as such usable by Statistics Finland, and vice versa. The missing interoperability complicates collaboration efforts, such as having a common data catalogue.

Statistics Finland uses the Common Structure of Statistical Information (CoSSI) data model, specifically developed for its data production needs. CoSSI supports the description of, for example, statistical publications, tables and datasets. FSD, on the other hand, utilises the international DDI data model, which is widely used by data archives to describe, especially, survey-type research materials. The development of the DDI model has also increased interest at statistical agencies.

The existing metadata models were compared for the purpose of finding a common language, creating interoperable metadata and enhancing cooperation possibilities. The comparison was carried out in the spring of 2014 as a collaboration between the metadata and researcher services of Statistics Finland, and FSD. The cooperation project further aimed at producing and describing public use files (PUFs) and enhancing description know-how in both organisations, especially regarding the DDI at Statistics Finland. This is necessary as the DDI data model is becoming increasingly popular in statistical organisations, and also Eurostat is examining it as a potential tool for describing microdata.

At FSD, the project was coordinated by Mari Kleemola, with Katja Moilanen participating in model comparison. At Statistics Finland, Saija Ylönen and Jussi Heino acted as project coordinators, while Mikko Saloila was in charge of data model comparison and conversion, and presenting these findings in the final report. The PUF metadata was written and described in the report by Annu Nissinen, and edited for test conversion by Satu Heinonen. Jukka Hämäläinen saw to the conversion application.

2 Metadata models: CoSSI and DDI Codebook

2.1 CoSSI

Statistics Finland's XML-based Common Structure of Statistical Information (CoSSI) is a modular model used for defining the basic structure and content of statistical information. It is used to describe, for example, statistical tables, classifications, concepts, variables, general information on statistical documents and quality descriptions. CoSSI was designed in accordance with international standards such as the Dublin Core and CALS. If needed, CoSSI can be expanded; new elements, e.g. for data descriptions, have already been integrated into it.

Preliminary examinations indicate that the CoSSI data model offers an adequate basis for producing content description data of statistical information following the GSIM (Generic Statistical Information Model) data model.¹

2.2 DDI Codebook

The DDI (Data Documentation Initiative) Codebook used by FSD is an effort to create an international standard for describing quantitative research data. Each dataset is described as precisely as possible, including, among other things, author(s), topic and content of the study, selection or the sampling method used, data collection, unit of observation/data unit, terms of use, file format(s), variables and the number thereof, question texts, and other metadata documents.

The DDI format is maintained and developed by the international DDI Alliance hosted by the University of Michigan. The members of the Alliance come from around the world and include data archives, universities, statistical offices, and organisations specialised in developing statistical and data collection software. DDI has been branched into two separate development lines:

- DDI Codebook (DDI 2), intended primarily for documenting and holding simple survey data; and
- DDI Lifecycle (DDI 3), offering solutions for periodical, reference and panel data documentation, and managing data across the entire life cycle.

DDI metadata are structured XML files. All FSD datasets are described using the DDI Codebook format and include about a third of its content elements.²

3 Data model comparison

3.1 CoSSI and DDI Codebook 2.1

The Social Science Data Archive (FSD) currently uses DDI Codebook version 2.1. DDI Codebook was developed specifically from the viewpoint of data archives and thus contains a lot of data archiving-related elements not present in CoSSI. On the other hand, DDI Codebook provides no support for time series essential for statistical agencies, or for classifications.

The data models are also structurally different. CoSSI is composed of modules, which can be used to describe different types of information. DDI Codebook, in turn, is built into a single entity with all parts contained under one root element.

A theoretical comparison of the data models was not crucial to the comparison table, but it did provide necessary background information for the project. The compilation of the table itself got a boost when the focus shifted to practice, i.e. to imple-

¹ Further information: http://www.stat.fi/org/tut/dthemes/drafts/cossi_en.html, <http://www1.unec.org/stat/platform/pages/viewpage.action?pagelid=78677992>

² Additional information <http://www.fsd.uta.fi/fi/aisteistot/arkistointi/ddi.html>, <http://www.ddialliance.org>

mentation between Statistics Finland's Variable Editor and FSD's DDI 2.1.

The comparison between the different general data elements (CoSSI's docmeta, DDI's document description) was considerably facilitated by the fact that both the data models are traceable to the Dublin Core data model through documentation. This made it easier to compare elements which had different names but basically meant the same.

At first, the work was slowed down by the fragmentation of the documentation, a trait shared by both CoSSI and DDI 2.1. Compared to CoSSI, DDI is a much more extensive. Statistics Finland utilises CoSSI's metadata elements almost in their entirety, while FSD uses only about a third of the fields included in the DDI Codebook format. A universally valid comparison of the models was, therefore, based on both FSD's own help documents and the universal tag library compiled by DDI Alliance.

The biggest differences in the data models were detected in technical data and single variable descriptions. This was expected, as the technical description of individual variables has not been FSD's key focus area. On the other hand, some of CoSSI's technical variable description elements are intended solely for Statistics Finland's own use, and provide no additional value for external users.

From the user's (usually a researcher's) viewpoint, the most significant data elements are those related to the description of the contents of the data, not those associated with technical aspects. Therefore, the gaps discovered in the data models pertaining to the said elements do not give rise to major concern.

The comparison table between CoSSI and DDI Codebook 2.1 is provided in Annex 1.

3.2 CoSSI and DDI Codebook 2.5

After DDI Codebook 2.1, it was decided to also include the upgraded version, 2.5, in the comparison. The aim was to examine changes brought forth by the new version in regard to conversion. The comparison between the old and the new version was further validated by the fact that FSD is in the process of transitioning to a new metadata system, and with it, to DDI Codebook version 2.5.

DDI Codebook 2.5 is primarily a bridge between DDI Lifecycle (3.1) and DDI Codebook that allows building mutually compatible models and, consequently, transporting data. After careful scrutiny, the new DDI Codebook 2.5 was found to differ from its predecessor in only three respects.

Comparison table regarding these (three elements) is appended as Annex 2.

3.3 DDI Lifecycle 3.1

Originally, the project intended to focus only on the comparison between CoSSI and DDI Codebook, but since the time-frame allowed it, DDI Lifecycle version 3.1 was briefly included in the study. Version 3.1 expands the DDI data model towards time series and classification descriptions. Of the different statistical agencies, DDI Lifecycle is primarily used in Australia and New Zealand.

DDI Lifecycle is considerably more extensive than DDI Codebook, and partly for this reason, difficult to manage. In addition, the documentation is still incomplete. At this point, one should keep an active eye on the development of DDI Life-

cycle. The Australian Bureau of Statistics has shouldered more responsibility within the Alliance to develop the specification, paving the way for better documentation and new versions.

3.4 Metadata test conversion

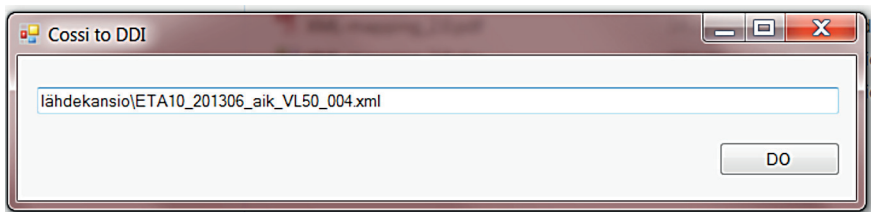
The comparison table was tested with a concrete XML conversion to verify the usability of the table.

For the purpose of conversion, a XML metadata file was constructed with Statistics Finland's variable editor, and then paired up with FSD's machine-readable metadata file³. On the basis of these files, an attempt was made to transport all information contained in Statistics Finland's metadata to the DDI-based structure. Some fields could not be transferred, in which case the data was dropped from the conversion. The element-level compatibility information is shown in the annexed conversion table.

A meeting with FSD led to some changes in the conversion: a decision was taken to add four additional elements, absent from the original test metadata. However, as these elements are included in the descriptions on a case by case basis, they had to be suitably incorporated into the DDI-based document.

The conversion application between CoSSI and DDI was written with C# programming language. The interface was designed to be as simple and easy as possible, containing only a field in which the CoSSI-based XML file that is to be converted is dragged.

After clicking the DO button, the DDI format XML file appears in the source folder of the original file. The name of the file shows in the following format: [original file name]_FSD.



Picture of the conversion application.

When examining the theoretical, i.e. documentation-based, and practical comparison table, a few exceptions were discovered on the basis of which the comparison table between CoSSI and DDI Codebook 2.1 was complemented.

The conversion table can be found in Annex 3.

³ The FSD's XML file used in the conversion was FSD1260, a description for Finnish National Election Study 2003: <http://www.fsd.uta.fi/fi/aaineistot/luettelo/FSD1260/meF1260.xml>

4. Public use files (PUFs)

4.1 Teaching datasets

Teaching datasets are public-use files especially suited for teaching purposes. Because the data are available to all, the privacy of the respondents needs to be particularly closely controlled. The data may not reveal the identity or characteristics of the study subjects, for which purpose the variables and observations contained in the material must be modified to avoid identification.

Teaching datasets are not suitable for conducting research. In teaching, correct research results are not of paramount importance as long as the dataset is structurally similar to the original one. Variables must, therefore, be similar or even have the same names as in the original data, and the distribution of the key background and study variables maintained. The results of the statistical analysis, however, do not necessarily correspond to those yielded by the original data.

During the present project, Statistics Finland produced two public-use teaching datasets: the time-use data PUF and the census PUF. The publication of the materials did not fall into the scope of this project. However, Statistics Finland published these two teaching datasets on its website in the autumn of 2014⁴. By the end of 2015, Statistics Finland aims to produce and publish a total of five public use files.

4.2 Time-use data PUF

In December 2013, the Turku Center for Welfare Research (TCWR) organised a seminar on time-use preferences that prompted the compilation of a PUF suitable for educational purposes.

The time-use PUF is a sample dataset, which makes it easier to employ disclosure avoidance techniques. The material included about 1400 person- and household-specific variables. Data protection measures included removing all non-key variables, leaving around 900 variables that were then rounded and blurred, inter alia, by aggregating them into sum variables, coarsening classifications, and interchanging different observation values. The teaching dataset cannot be linked with ID variables to the original research data. During data modification, certain background variable values were also changed so that the observations included in the teaching dataset no longer correspond with actual study participants.

4.3 Census PUF

In Finland, population census and the related count of dwellings and real estate are carried out every ten years in accordance with the Population Census Act. The general, direct censuses span from 1950 to the present, making the Finnish census data internationally unique.

The PUF dataset intended for teaching purposes focuses on the 0-9-year-olds in the 1950 census data, i.e. the so-called baby boomers, of whom 1% were randomly sampled.

⁴ <http://www.stat.fi/tup/mikroaineistot/opetusaineistot.html> (in Finnish only)

The PUF contains information on the sampled individuals from the years 1950, 1970, 1980, 1990, 2000 and 2010. The dataset covers 16 variables, all of which were included in all the censuses. Variables have been largely harmonised for better comparison. For teaching purposes, variables tapping into the line of business and socioeconomic status were left in the dataset. These types of variables, where classifications vary over the years, provide an excellent tool for learning.

Data has been retrieved for the same individuals sampled from the original 1950 census, which means that the number of people gradually decreases. In a 1% random sample, the risk of identification is very small. The dataset does not contain sensitive information, and area data, for example, are provided only at the level of major regions. A municipality group variable is included, enabling the examination of, for example, migration from the countryside to urban areas. Each sampled person has been assigned an individual ID number, which cannot be linked to other Statistics Finland resources.

In addition to the above-mentioned variables, the dataset contains a family identification code through which the data of the children can be attached to a separate parents file. The parents file will contain information corresponding to the PUF variables gathered from the 1950 census data.

5. Publication of metadata in Statistics Finland's web service

At the moment, Statistics Finland's ready-made metadata is available at <http://tilastokeskus.fi/tup/mikroaineistot/aineistot.html>

The plan is to publish them in Statistics Finland's web-based data catalogue, which would also serve as a publication forum for metadata compiled by other authorities. The catalogue will be coupled with search engines that allow easy access to desired variables, datasets and topic areas.

The data catalogue will be released in 2014-15 as part of Statistics Finland's web service reform, and is funded through the Finnish Microdata Access Services (FMAS⁵) project.

6. Summary

The primary objective of the project was to compare the different data models used by the two organisations, and to test their compatibility. In addition, Statistics Finland produced two public use files (PUFs) with metadata.

The comparison between CoSSI and DDI Codebook data models provided new information regarding the compatibility of the models. Element-specific comparison tables have been annexed to this report.

The biggest differences in the data models pertained to technical data and single variable descriptions. This was expected, as the technical description of individual variables has not been FSD's key focus area. On the other hand, some of CoSSI's

⁵ For more information on FMAS: <http://fmasfoorumi.wordpress.com/>

technical variable description elements are intended solely for Statistics Finland's own use, and provide no additional value for external users.

Comparison between the general data elements was made easier by the fact that both models are traceable through documentation to the Dublin Core data model. From the viewpoint of a dataset catalogue serving the needs of researchers, the most important elements relate to content, while detailed technical data descriptions are often secondary. Thus, the gaps discovered in the elements describing the technical aspects of the data should not cause major problems.

The comparison found the CoSSI documentation to be incomplete. The current metadata document is a XML document with a CSS style sheet. During the project, a proposal was put forward to improve the readability of the documentation especially to a person with limited XML knowledge. The proposal focused on better management of CoSSI's structure. The DDI documentation was also perceived as difficult to navigate and handle.

The project also produced PUF dataset versions and variable-level descriptions to be released later on Statistics Finland's website. Statistics Finland is building a new material catalogue as part of its website reform.

In all, this project deepened metadata expertise at both Statistics Finland and FSD, and significantly increased cooperation between the two organisations in developing metadata and researcher services. Ways to enhance cooperation even further were identified, including, for example:

- technical solutions for data description and metadata management
- study/register/statistics availability from researchers' viewpoint
- marketing and communication to dataset users and potential users
- national solutions related to dataset and metadata distribution.

7. Additional information

CoSSI data model:

- http://www.stat.fi/org/tut/dthemes/drafts/cossi_en.html
- Saija Ylönen, Statistics Finland. [saija.ylonen \[a\] tilastokeskus.fi](mailto:saija.ylonen@tilastokeskus.fi)

DDI data model:

- <http://www.ddialliance.org/>

PUF datasets:

- Jussi Heino, Statistics Finland. [jussi heino \[a\] tilastokeskus.fi](mailto:jussi.heino@tilastokeskus.fi)

Statistics Finland's metadata:

- Saija Ylönen, Statistics Finland. [saija.ylonen \[a\] tilastokeskus.fi](mailto:saija.ylonen@tilastokeskus.fi)

FSD's DDI model:

- Mari Kleemola, FSD. [mari.kleemola \[a\] uta.fi](mailto:mari.kleemola@uta.fi)

Annex 1

Comparison table CoSSI–DDI Codebook (2.1) and Variable editor's data field descriptions

Mapping ID#	The data field name in the variable editor	CoSSI	DDI2.1	DDI2.1 identification
		Docmeta	docDscr* / stdyDscr+	1.0 / 2.0
1	Content language	main_language, other_language*	codebook (att xml:lang)	0.0
2	Created	day, time	prodDate* (no field for documenting specific time)	1.1.3.3
3	Modified	day, time	version? (att date) (no field for documenting specific time)	1.1.6.1
4	Subject field	categories (att type)	topcClas*	2.2.1.2
5	Division	seriename	serName*	2.1.5.1
6	Maintenance	maintenance	frequenc*	2.3.1.3
7	Reference time	coveragetemporal*	timePrd* (att cycle, start, end, single)	2.2.3.1
8	Subject	subject	titl	2.1.1.1
9	Population	coveragepopulation*	universe*	2.2.3.9
10	Keywords	keywords	keyword*	2.2.1.1
11	Data source	resourcerelation*	dataSrc*	2.3.1.8.1
12	Content description	contentdescription	abstract*	2.2.2
13	creator	surname, first_name	AuthEnty*	1.1.2.1
14	contributor	surname, first_name	othId*	1.1.2.2
15	First name	first_name	-	
16	Surname	surname	-	
17	Phonenumber	phonenumber	-	
18	email_fi/sv/en	email	-	
19	hard coded in the variable editor	orgname	producer*	1.1.3.1
		address, wwwaddress etc.	-	
20	hard coded in the variable editor	type	-	
21	hard coded in the variable editor	(URN URL ISBN ISSN handle stojid)*	IDNo* (att agency)	1.1.1.5
22	hard coded in the variable editor	rights*	copyright?	1.1.3.2
		Statmeta	dataDscr	4.0
23	Short name	variable	var* (att ID)	4.3
24	Long name	variablename	labl*	4.3.2
25	Concept definition	conceptdef	concept	4.3.21
26	Measure	measunit	var* (att measUnit)	4.3
27	Operational definition	operdef	No equivalent element in DDI2. The element notes* can be used.	4.3.26
28	Deductionrule	deductionrule	No equivalent element in DDI2. The element notes* can be used.	4.3.26
29	Variable modified_date	varmodifieddate	version? (att date)	4.3.20.1
30	Variable startdate	varstartdate	-	
31	Variable enddate	varenddate	-	
32	Variable group	variablegroup	varGrp*	4.1
33	Workcomment	workcomment	notes*	4.3.26
34	VarType	vartype	varFormat? (att type)	4.3.23
35	Min_value	valuerangemin	sumStat* (att type="min")	4.3.14
36	Max_value	valuerangemax	sumStat* (att type="max")	4.3.14
37	Valuerange list	valuerangelist	valrng*	4.3.9
38	Lenght	recordlength	location* (att width)	4.3.1
39	Starting position	startingposition	location* (att Startpos)	4.3.1
40	Missing value codes	missingvalue	invalrng*	4.3.10
41	Checking macros	checkingmacros	drvcmd?	4.3.22.2

Annex 1 ...

Mapping ID#	The data field name in the variable editor	CoSSI	DDI2.1	DDI2.1 identification
		filemeta (procmeta)		
42	Filename	filename	fileName? (att ID)	3.1.1
43	Fileformat	fileformat	fileType?	3.1.5
44	Filecomment	filecomment	notes*	3.1.3.2
	No equivalence to DDI			
45	Variable old name	varoldname	-	
46	Classification_id	classification (att classificationid)	-	
47	Status	variable (att status)	-	
48	Format	varformat	-	
49	Missing values (allowed)	missingvalues (att allowed)	-	
50	Indexed var	indexedvar (att type)	-	
51	Primary key	att primarykey	-	
52	Foreign key	att foreignkey	-	
	Not used in the variable editor			
	The variable specification element is used when the variable name is not descriptive enough. The element contains a more detailed description of the variable.	variablespecification	txt or label	4.3.15 or 4.3.2
	Description of the operational definition	operdefdesc	-	
	The element contains calculation rules.	calculation	-	
	If needed, used to specify measurement unit.	measdesc	-	
	Name of a figure file attached to the variable	figuregrp	-	
	Not used in the variable element (docmeta)			
	Date: Finished (finalised)	day	-	
	The format is either XML or XHTML (used for e-publications)	format	-	
		categories (att type)	topcClas*	2.2.1.2
		(URN URL ISBN ISSN handle stojid)*	IDNo* (att agency)	1.1.1.5
		coveragespatial*	geogCover*	2.2.3.4
	Equivalent to Dublic Core source element	source	dataSrc*	2.3.1.8.1

Annex 2
Comparison table CoSSI–DDI Codebook 2.5 (as per changed elements)

Mapping ID#	The data field name in the variable editor	CoSSI	DDI2.1	DDI identification	DDI 2.5
		Docmeta	docDscr* / stdyDscr+	1.0 / 2.0	
2	Created	day, time	prodDate* (no field for documenting specific time)	1.1.3.3	New date fields follow the ISO format (YYYY, YYYY-MM, and YYYY-MM-DD plus time.) 2010-10-14T14:42:23
9	Population	coveragepopulation*	universe*	2.2.3.9	New structure (Sample frame). More metadata on sampling. Note: DDI id may change.
11	Data source	resourcerelation*	dataSrc*	2.3.1.8.1	New element to specify an external source (sourceCitation).

Annex 3

Comparison table for XML conversion

Comparison table for XML conversion			
#	Cossi-XML	DDI-XML	Note
1	<first_name>Pihla</first_name>	<docDscr> <rspStmt> Merimaa, Pihla </AuthEnty> Used AuthEnty	<citation> <AuthEnty> No specific elements for contact information. All such information documented in the same element.
2	<surname>Merimaa</surname>		
3	<phonenumber>+358 9 1734 3508</phonenumber>		Not used at FSD, dropped.
4	<email xml:lang="fi">pihla.merimaa@stat.fi</email>		Not used at FSD, dropped.
5	<subject xml:lang="fi">VL50-otosaineiston henkilöt väestölaskentavuosina vuoteen 2010 (PUF) </subject>	<styDscr> <citation> <titl></titl>	
6	<contentdescription xml:lang="fi">VL50-otosaineistosta on poimittu 1% satunnaisotos 0-9 -vuotiaista, eli suuriin ikäluokkiin kuuluvista henkilöistä. Henkilöille on poimittu tietoja vuosilta 1950, 1970, 1980, 1990, 2000 Muuttujat on yhdenmukaistettu, poislukien sosioekonominen asema ja toimialatiedot ja 2010.	<styDscr> <styInfo> <abstract> </abstract>	
7	<orgname>Tilastokeskus</orgname>	<docDscr> <prodStmt> <producer> abbr="TK">Tilastokeskus</producer>	<citation> No specific elements for contact information. All such information documented in the same element.
8	<phonenumber>17341</phonenumber>		Not used at FSD, dropped.
9	<address>Työpajankatu 13</address>		Not used at FSD, dropped.
10	<wwwaddress>www.tilastokeskus.fi</wwwaddress>		Not used at FSD, dropped.
11	<fax>1734 2279</fax>		Not used at FSD, dropped.
12	<first_name>Satu</first_name>	<docDscr> <rspStmt> <othid> Heinonen, Satu </othid> Used AuthEnty	<citation> No specific elements for email addresses or phone numbers.
13	<surname>Heinonen</surname>		
14	<email xml:lang="fi">satu.heinonen@stat.fi</email>		Not used at FSD, dropped.
15	<phonenumber>+358 9 1734 2751</phonenumber>		Not used at FSD, dropped.
16	<day>2013-07-09</day>	<docDscr> <prodStmt> <prodDate date="2005-01-20"/>	<citation>
17	<time>15:29:16</time>		No element for documenting specific time in DDI.
18	<modified><day>2013-12-17</day>	<docDscr> <verStmt> <version date="2012-01-05"> </version>	<citation>
19	<time>10:25:22</time>		No element for documenting specific time in DDI.
20	<typegrp> <type xml:lang="fi"> Tilastollinen metatieto</type>		Not used at FSD, dropped.
21	<main language>fi</main language>	<codeBook xml:lang="fi">	

22	<serienam xml:lang="fi">Väestölaskennat</serienam>	<stdyDscr> <citation> <serStmt> <serName >Väestölaskennat </serName> <stdyDscr> <method> <dataColl> <frequenc>1/10A </frequenc> <stdyDscr> <stdyInfo> <subject> <topcClas>VRM</topcClas> </subject> <docDscr> <citation> <IDNo> vaelaskp_2013-07_2013-07-09_ain_0001 </IDNo> <docDscr> <copyright>Tilastokeskus</copyright> <prodStmt> <stdyDscr> <stdyInfo> <universe clusion="1"></universe> <stdyDscr> <method> <dataColl> <sources> <dataSrc> Väestölaskennat 1950, 1970; 1980, 1990, 2000 ja 2010 </dataSrc> <dataDscr> Name="vaelaskp_2013-07_2013-11-05_muu_0001" ID="vaelaskp_2013-07_2013-11-05_muu_0001" > <dataDscr> <var> <labl level="variable">[vaelaskp_2013-07_2013-11-05_muu_0001] Henkilön tunnistus </labl> <dataDscr> <var> <concept>Henkilön tunnistenumero aineistossa </concept> </concept>	
23	<maintenance xml:lang="fi">1/10A</maintenance>		
24	<categories type="vrm"/>		
25	<DOI>vaelaskp_2013-07_2013-07-09_ain_0001</DOI>		
26	<rights xml:lang="fi">Tilastokeskus</rights>		
27	<coveragepopulation xml:lang="fi">VL50:n henkilöt</coveragepopulation>		
28	<resource relation xml:lang="fi">Väestölaskennat 1950, 1970; 1980, 1990, 2000 ja 2010</resource relation>		Not used at FSD, dropped. The attribute "status" dropped.
29	<variable fieldName="id" status="processed" variableId="vaelaskp_2013-07_2013-11-05_muu_0001">		
30	<variableName xml:lang="fi">Henkilön tunnistus</variableName>		
31	<conceptDef xml:lang="fi">Henkilön tunnistenumero aineistossa</conceptDef>		
32	<varmodifieddate>2013-11-04</varmodifieddate>		Not used at FSD, dropped.
33	<missingValues allowed="no"/>		Not used at FSD, dropped.
34	<varType varType="numeric"/>		Not used at FSD, dropped.
35	<varformat>int</varformat>		Not used at FSD, dropped.

Annex 3 ...

36	<deductionrule>Perustuu 1950 väestötlaskennan lomaketietoihin.</deductionrule>	<dataDscr> <var> <notes>Perustuu 1950 väestötlaskennan lomaketietoihin. </notes>	
37	<operdef xml:lang="fi">Tieto saatavilla 1950, 1970, 1980, 1990, 2000 ja 2010.</operdef>	<dataDscr> <var> <notes>Total population for the agency for the year reported. </notes>	
38	<valuerange> <valuerangemin>1</valuerangemin> <valuerangemax>2</valuerangemax> </valuerange>	<dataDscr> <var> <valmg> <range max="2" min="1"/></valmg>	
39	<recordlength>1</recordlength>		Not used at FSD, dropped.
40	<measunit xml:lang="fi">lkm</measunit>	<dataDscr> measUnit="lkm"> <var>	MeasUnit is a attribute of the element "var" .
41	<measdesc xml:lang="fi"> Lukumäärä</measdesc>		Not used at FSD, dropped.
42	<valuerange> <valuerangelist delimiter=" " ">1-5,9</valuerangelist> </valuerange>	<dataDscr> <var> <valmg><item VALUE="1" /><item VALUE="2" /><item VALUE="3" /></valmg>	
43	<keywords> </keywords>	<stdyDscr> <subject> <keyword> </keyword> <stdyInfo>	
44	<coveragetime> </coveragetime>	<stdyDscr> <sumDscr> <timePrd event="start" date="YYYY-MM-DD"> </timePrd> event="end" date="YYYY-MM-DD"> </timePrd>	
45	<filename> </filename>	<fileDscr> <fileTxt> <fileName ID="XX">Tiedoston nimi </fileName> <fileDscr> <fileTxt> </format>	The file is given an ID attribute, and variables contain a reference to it.
46	<fileformat> </fileformat>		



FINNISH SOCIAL
SCIENCE DATA ARCHIVE

Visiting address: Åkerlundinkatu 5 B, 4th floor
Tampere, Finland

Postal address: Finnish Social Science Data Archive
FI-33014 University of Tampere, FINLAND

Email: fsd@uta.fi

Internet: www.fsd.uta.fi

